

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2002, 00:09:11 ; Search time 8498.8 Seconds

(without alignments)
31.610 Million cell updates/sec

Title: US-09-851-670-4

Perfect score: 25

Sequence: 1 acagtagcagcacagcatgagacc 25

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues 111874

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 60

Post-processing:

Maximum Match 0%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_estlin:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_estlin:*
9: em_hic:*
10: gb_estl:*
11: gb_estl:*
12: gb_hic:*
13: gb_hic:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_pro:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	16.2	64.8	58	10	A1759324
2	16	64.0	32	11	N40655
3	16	64.0	55	11	BF346180
4	15.2	60.8	52	10	AA996148
5	15	60.0	38	13	AA002540
6	15	60.0	49	11	H66141
7	14.8	59.2	52	11	C20867
8	14.8	59.2	53	13	TA180E05P
9	14.8	59.2	57	13	A2616824
10	14.6	58.4	24	13	A2779573
11	14.6	58.4	27	13	A2404206
12	14.6	58.4	35	13	A2861400

C 13	14.4	57.6	40	13	A2832139	A2832139
C 14	14.4	57.6	51	13	A2488023	A2488023
C 15	14.4	57.6	53	13	TA263C05P	TA263C05P
C 16	14.4	56.0	46	10	A1884025	A1884025
C 17	14.4	56.0	50	10	AU103544	AU103544
C 18	14.4	56.0	50	10	AU104442	AU104442
C 19	14.4	56.0	52	10	AA856040	AA856040
C 20	14.4	56.0	52	10	AA615079	AA615079
C 21	13.8	55.2	50	10	AU102281	AU102281
C 22	13.8	55.2	50	10	AU105836	AU105836
C 23	13.8	55.2	57	10	BE408921	BE408921
C 24	13.6	54.4	28	10	AU1755903	AU1755903
C 25	13.6	54.4	56	10	AA587506	AA587506
C 26	13.6	54.4	60	13	A2651227	A2651227
C 27	13.4	53.6	50	10	AU102654	AU102654
C 28	13.4	53.6	52	10	BE321070	BE321070
C 29	13.4	53.6	56	13	A2820546	A2820546
C 30	13.2	52.8	21	13	A2831993	A2831993
C 31	13.2	52.8	38	13	A2971376	A2971376
C 32	13.2	52.8	39	13	A2635338	A2635338
C 33	13.2	52.8	44	10	AV836720	AV836720
C 34	13.2	52.8	50	10	AV833524	AV833524
C 35	13.2	52.0	32	13	TA337F07Q	TA337F07Q
C 36	13.2	52.0	37	13	A2807121	A2807121
C 37	13.2	52.0	41	10	AA509356	AA509356
C 38	13.2	52.0	43	10	AA937113	AA937113
C 39	13.2	52.0	50	11	BE537767	BE537767
C 40	13.2	52.0	53	13	AF179999	AF179999
C 41	13.2	52.0	53	13	AF180000	AF180000
C 42	13.2	52.0	54	11	BF163806	BF163806
C 43	13.2	52.0	56	13	HSWC02806	HSWC02806
C 44	13.2	52.0	59	10	AM687832	AM687832
C 45	13.2	52.0	60	13	A2623437	A2623437

ALIGNMENTS

RESULT 1
A1759324
LOCUS E1759324 58 bp mRNA EST 18-JAN-2000
DEFINITION E1759324.1 GI:5174991
ACCESSION A1759324
VERSION A1759324.1 GI:5174991
KEYWORDS
SOURCE
ORGANISM
Eimeria tenella
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;

REFERENCE
AUTHORS
Liberator, P., Diaz, C., Tang, R., Marra, M., Hillier, L., Kucaba, T., Martin, J., Wylie, T., Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Florence, N., Shih, T., Jackson, T., Cardenas, M., McCann, R., Waterson, R., Wilson, R. and Sibley, D.
WashU-Merck Eimeria tenella project
Unpublished (1999)
Contact: David Sibley, Ph.D.
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Contact David Sibley (toxest@wustl.edu) for further
information relating to organism, libraries, or clone availability.
Trace considered overall poor quality
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 58
/organism="Eimeria tenella"

FEATURES
source

RESULT 4
AA96148 LOCUS
DEFINITION AA96148 52 bp mRNA EST 13-APR-1999
os14d11.s1 NCI_CGAP_Lu5 Homo sapiens CDNA clone IMAGE:1605309 3'
similar to SW:ASH1_HUMAN P50553 ACHAEFE-SCUTE HOMOLOG 1.; mRNA
sequence.
ACCESSION AA96148
VERSION AA96148.1 GI:3182637
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 52)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-femail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES
source
Trace considered overall poor quality
Insert Length: 753 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1605309"
/clone_lib="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pRT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand CDNA was prepared from
neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded CDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pRT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 12 a 19 c 16 g 5 t
ORIGIN

Query Match 60.8%; Score 15.2; DB 10; Length 52;
Best Local Similarity 85.0%; Pred. No. 2.5e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 cagtagcagcacacagcatga 21
||| ||||| ||||| |
DB 5 CAGCAGCAGCAGCAGCATCA 24

RESULT 5
AA025540 LOCUS
DEFINITION AA025540 38 bp DNA GSS 23-AUG-2000
EP(X)1614-5prime Drosophila melanogaster EP line Drosophila
melanogaster genomic Sequence recovered from 5' end of P element,
DNA sequence.
ACCESSION AA025540
VERSION AA025540.1 GI:3265892
KEYWORDS GSS.

SOURCE
ORGANISM Drosophila melanogaster
fruit fly.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 38)
Liao,G.-C., Rehm,E.J. and Rubin,G.M.
Insertion site preferences of the P transposable element in
Drosophila melanogaster.
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
20202638
JOURNAL Contact: Gerald Rubin
MEDLINE Berkeley Drosophila Genome Project
UNIVERSITY University of California, Berkeley
COMMENT USA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu
Sequence recovery method was inverse PCR.
Sequence orientation is forward strand relative to 5' end of P
element

FEATURES
source
The P element insertion position is base 31 in the 38 bases. This
insertion position refers to the first base of the 8 base target
recognition sequence.
Class: transposon-tagged.
Location/Qualifiers
1..38
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster EP line"
/note="Inverse PCR was performed on Drosophila
melanogaster strains each of which contains a single EP
transposable element insertion. (The generation of these
insertion strains is described in Roth P, Szabo K, Bailey
A, Laverly T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
V, Ansoerg M, Cohen SM. Development 6:1049-1057.) The
resultant fragment for each strain was directly sequenced
to determine the genomic sequence at the site of
insertion. Details of the protocols used can be found at
http://fruitfly.berkeley.edu/P_p-disrupt/inverse_pcr.html."

BASE COUNT 14 a 7 c 11 g 6 t
ORIGIN

Query Match 60.0%; Score 15; DB 13; Length 38;
Best Local Similarity 78.3%; Pred. No. 2.9e+04;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 acagttagcagcacacagcatga 23
||||| ||| | ||| |||
DB 9 ACAGTATCAGCAGCAGATTGAGA 31

RESULT 6
H66141/c LOCUS
DEFINITION H66141 49 bp mRNA EST 18-OCT-1995
yu16e05.s1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone
IMAGE:233984 3' similar to gb|U87903|HUMAN|NE36 Human carcinoma
cell-derived Alu RNA transcript. (rRNA): gb:M55531 GALACTOSIDE
2-L-FUCOSYLTRANSFERASE (HUMAN); mRNA sequence.
ACCESSION H66141
VERSION H66141.1 GI:1024881
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 49)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston

Db 40 AGCAGCAGCAGCAGCAGA 23

||||| ||||| |||||

RESULT 9

A2616824 57 bp DNA

GSS 13-DEC-2000

LOCUS 1M0446L15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0446L15 R, DNA sequence.

ACCESSION

A2616824

GSS

VERSION

A2616824.1 GI:11739014

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE

1 (bases 1 to 57)

JOURNAL

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

COMMENT

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

FEATURES

,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.

SOURCE

and Wright,D., Weiss,R.

ORGANISM

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE

plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

FEATURES

University of Utah

SOURCE

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT

TITLE

84112, USA

JOURNAL

Tel: 801 585 5606

COMMENT

Fax: 801 585 7177

FEATURES

Email: ddunne@genetics.utah.edu

SOURCE

Insert Length: 10000 Std Error: 0.00

TITLE

Plate: 0446 row: L column: 15

JOURNAL

Seq primer: CACACGAGAACACGCTATGACC

COMMENT

Class: plasmid ends

FEATURES

High quality sequence stop: 57.

SOURCE

Location/Qualifiers

TITLE

1..57

JOURNAL

/organism="Mus musculus"

COMMENT

/strain="C57BL/6J"

FEATURES

/db_xref="taxon:10090"

SOURCE

/clone="UUGC1M0446L15"

TITLE

/clone.lib="Mouse 10kb plasmid UUGC1M library"

JOURNAL

/sex="Male"

COMMENT

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

FEATURES

/note="Vector: PMD42nv; Purified genomic DNA from M.

SOURCE

laboratory Mouse DNA Resource

TITLE

(http://www.jax.org/resources/documents/dnares/). The DNA

JOURNAL

was hydrodynamically sheared by repeated passage through a

COMMENT

0.005 inch orifice at constant velocity. The sheared DNA

FEATURES

was blunt end-repaired with T4 DNA polymerase and T4

SOURCE

polynucleotide kinase. Adaptor oligonucleotides were

TITLE

ligated to the blunt ends in high molar excess. The

JOURNAL

adapted DNA was purified and size-selected for a 9.5 to

COMMENT

10.5 kb range using preparative agarose gel

FEATURES

electrophoresis. Vector DNA was prepared from a derivative

SOURCE

of PMD42 (g11473211419b1AF129072.1), a copy-number

TITLE

inducible derivative of plasmid R1. The vector was ligated

JOURNAL

with adaptors complementary to the insert adaptors and

COMMENT

purified. The sheared, adapted mouse DNA was annealed to

FEATURES

adapted vector DNA, and transformed into

SOURCE

chemically-competent E. coli XL10-Gold (Stratagene) cells

TITLE

and selected for ampicillin resistance."

JOURNAL

BASE COUNT

0 a 7 c 25 g 25 t

COMMENT

ORIGIN

FEATURES

Query Match 59.2%; Score 14.8; DB 13; Length 57;

SOURCE

Best Local Similarity 88.9%; Pred. No. 3.6e+04;

TITLE

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acagtagcagcagcagca 18

||||| ||||| |||||

RESULT 10

A2779573

24 bp DNA

GSS 16-FEB-2001

LOCUS 2M0016K09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0016K09 F, DNA sequence.

ACCESSION

A2779573

VERSION

A2779573.1 GI:12910362

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE

1 (bases 1 to 24)

JOURNAL

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

COMMENT

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

FEATURES

,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.

SOURCE

and Wright,D., Weiss,R.

ORGANISM

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE

plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

FEATURES

University of Utah

SOURCE

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT

TITLE

84112, USA

JOURNAL

Tel: 801 585 5606

COMMENT

Fax: 801 585 7177

FEATURES

Email: ddunne@genetics.utah.edu

SOURCE

Insert Length: 10000 Std Error: 0.00

TITLE

Plate: 0016 row: K column: 09

JOURNAL

Seq primer: CGTGTAAACGACGCGCAGT

COMMENT

Class: plasmid ends

FEATURES

High quality sequence stop: 24.

SOURCE

Location/Qualifiers

TITLE

1..24

JOURNAL

/organism="Mus musculus"

COMMENT

/strain="C57BL/6J"

FEATURES

/db_xref="taxon:10090"

SOURCE

/clone="UUGC2M0016K09"

TITLE

/clone.lib="Mouse 10kb plasmid UUGC1M library"

JOURNAL

/sex="Male"

COMMENT

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

FEATURES

/note="Vector: PMD42nv; Purified genomic DNA from M.

SOURCE

musculus C57BL/6J (male) was obtained from the Jackson

TITLE

laboratory Mouse DNA Resource

JOURNAL

(http://www.jax.org/resources/documents/dnares/). The DNA

COMMENT

was hydrodynamically sheared by repeated passage through a

FEATURES

0.005 inch orifice at constant velocity. The sheared DNA

SOURCE

was blunt end-repaired with T4 DNA polymerase and T4

TITLE

polynucleotide kinase. Adaptor oligonucleotides were

JOURNAL

ligated to the blunt ends in high molar excess. The

COMMENT

adapted DNA was purified and size-selected for a 9.5 to

FEATURES

10.5 kb range using preparative agarose gel

SOURCE

electrophoresis. Vector DNA was prepared from a derivative

TITLE

of PMD42 (g11473211419b1AF129072.1), a copy-number

JOURNAL

inducible derivative of plasmid R1. The vector was ligated

COMMENT

with adaptors complementary to the insert adaptors and

FEATURES

purified. The sheared, adapted mouse DNA was annealed to

SOURCE

adapted vector DNA, and transformed into

TITLE

chemically-competent E. coli XL10-Gold (Stratagene) cells

JOURNAL

and selected for ampicillin resistance."

COMMENT

BASE COUNT

8 a 8 c 8 g 0 t

ORIGIN

Query Match 58.4%; Score 14.6; DB 13; Length 24;

FEATURES

Best Local Similarity 81.0%; Pred. No. 3.8e+04;

SOURCE

Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 cagtagcagcacacagcatgag 22
Db 2 CAGCAGCAGCAGCAGCAGCAG 22

RESULT 11
A2404206/c 27 bp DNA GSS 03-OCT-2000
LOCUS 1M017120F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M01712120 F, DNA sequence.
ACCESSION A2404206
VERSION A2404206.1 GI:10528219
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 27)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0172 row: 1 column: 20

Seq primer: CGTTCGTAACACGACGCCACT
Class: plasmid ends

High quality sequence stop: 27.

FEATURES
source Location/Qualifiers

1..27
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M01712120"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114|gblAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

0 a 9 c 9 g 9 t

Query Match 58.4%; Score 14.6; DB 13; Length 27;
Best Local Similarity 81.0%; Pred. No. 3.8e+04;

Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 2 cagtagcagcacacagcatgag 22
Db 27 CAGCAGCAGCAGCAGCAGCAG 7

RESULT 12
A2861400 35 bp DNA GSS 21-FEB-2001
LOCUS 2M0167013R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0167013 R, DNA sequence.
ACCESSION A2861400
VERSION A2861400.1 GI:13057682
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 35)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0167 row: 0 column: 13

Seq primer: CACACAGGAACACGATGACGC
Class: plasmid ends

High quality sequence stop: 35.

FEATURES
source Location/Qualifiers

1..35
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0167013"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114|gblAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

12 a 11 c 12 g 0 t

Query Match 58.4%; Score 14.6; DB 13; Length 35;

Query Match 57.6%; Score 14.4; DB 13; Length 51;
 Best Local Similarity 75.0%; Pred. No. 5e+04;
 Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 2 cagtagcagcaacagcatgagacc 25
 ||| | |||| | ||| | |||
 Db 45 CACGACCAGCAGCACCACATGACACC 22

RESULT 15

TA263C05P/c 53 bp DNA GSS 13-DEC-2000

LOCUS T. brucei sheared genomic DNA clone 263c05, forward sequence,

DEFINITION genomic survey sequence.

ACCESSION AL483802

VERSION AL483802.1 GI:11849892

KEYWORDS GSS.

SOURCE Trypanosoma brucei.

ORGANISM Trypanosoma brucei

Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

1 (bases 1 to 53)

Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,

Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,

Melville,S.E., Rajandream,M.A. and Barrell,B.G.

Direct Submission

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing

Project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nhs@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TRU927/4 GUTat 10.1) was mechanically sheared

to give a tight size distribution (

4 kb). The v + 1 method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In

Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available

at http://www.sanger.ac.uk/Projects/T_brucei/.

Location/Qualifiers

1. 53

/organism="Trypanosoma brucei"

/strain="TRU927"

/db_xref="taxon:5691"

/clone="263c05"

BASE COUNT 5 a 12 c 14 g 22 t

ORIGIN

Query Match 57.6%; Score 14.4; DB 13; Length 53;

Best Local Similarity 75.0%; Pred. No. 5e+04;

Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 2 cagtagcagcaacagcatgagacc 25

||| | |||| | ||| | |||

Db 36 CACAAGCAGACGACGACCTAGACC 13

Search completed: March 9, 2002, 00:09:14
 Job time: 11030 sec